

# FAQ Retrieval using Query-Question Similarity and BERT-Based Query-Answer Relevance

Wataru Sakata†, Tomohide Shibata‡, Ribeka Tanaka‡, Sadao Kurohashi‡  
 †LINE Corporation, ‡Kyoto University

Download our dataset and codes



## Issues

- Train the neural model using the relevance label between user's q and appropriate QA [Karan+18]  
 → **It costs a lot to construct such labeled data.**
- Using the relevance between a question and an answer trained by QA pairs in the community QA site such as Yahoo QA [Shafiq+18].  
 → **The size of QAs in FAQ is not often enough to train a model.**

## Proposed Retrieval Model

Query : Can you please remove the wasp's nest?

**q-Q similarity**

**q-A relevance**

- Due to the scarcity of the labeled data, we use the BM25-based unsupervised retrieval system, **TSUBAKI** [Shinzato+12]
- High precision but difficult to deal with lexical gap.
- Trained with QA data using **BERT** [Devlin+19].
- **Collecting other similar FAQ sets to increase the size of available QA data.**
- It performs well even if there is a lexical gap between q and Q.

**Target FAQ**

Q: Please tell me how to take out the oversized garbage?	A: Please call the center, or you can also bring it there by yourself.
Q: I want to get a tax certificate by mail.	A: For the mail of the certificate, ...
Q: I want you to exterminate the bee's nest.	A: When there are bee nests in residential buildings...
...	...

## Combining TSUBAKI and BERT

1. For QA pairs whose q-Q similarity gets a higher score than  $\alpha$ , we rank them in order of the similarity.
2. Then for others, we rank them in order of the weighted average of q-Q similarity and q-A relevance.

## How to use BERT?

Fine-tuning stage QA pairs collected from Japanese local government FAQ pages

Qa1: Tell me how to take out the large trash.	Aa1: Call the Living Environment Center and ...
Qa2: Can I get a copy of my family register by mail?	Aa2: Yes. To request the certificate, ...
...	...
Qb1: How to throw away oversized trash?	Ab1: Please call the center.
...	...
Qk1: Please tell me how to take out the oversized garbage.	Ak1: Please call the center or you can also bring it...
Qk2: I want to get a tax certificate by mail.	Ak2: For the mail of the certificate, ...
...	...

Search stage (for the city k)

q: I want to throw away a large waste.	0.92	Ak1: Please call the center or you can also bring it...
	0.02	Ak2: For the mail of the certificate, ...
	0.01	...

## Experiments

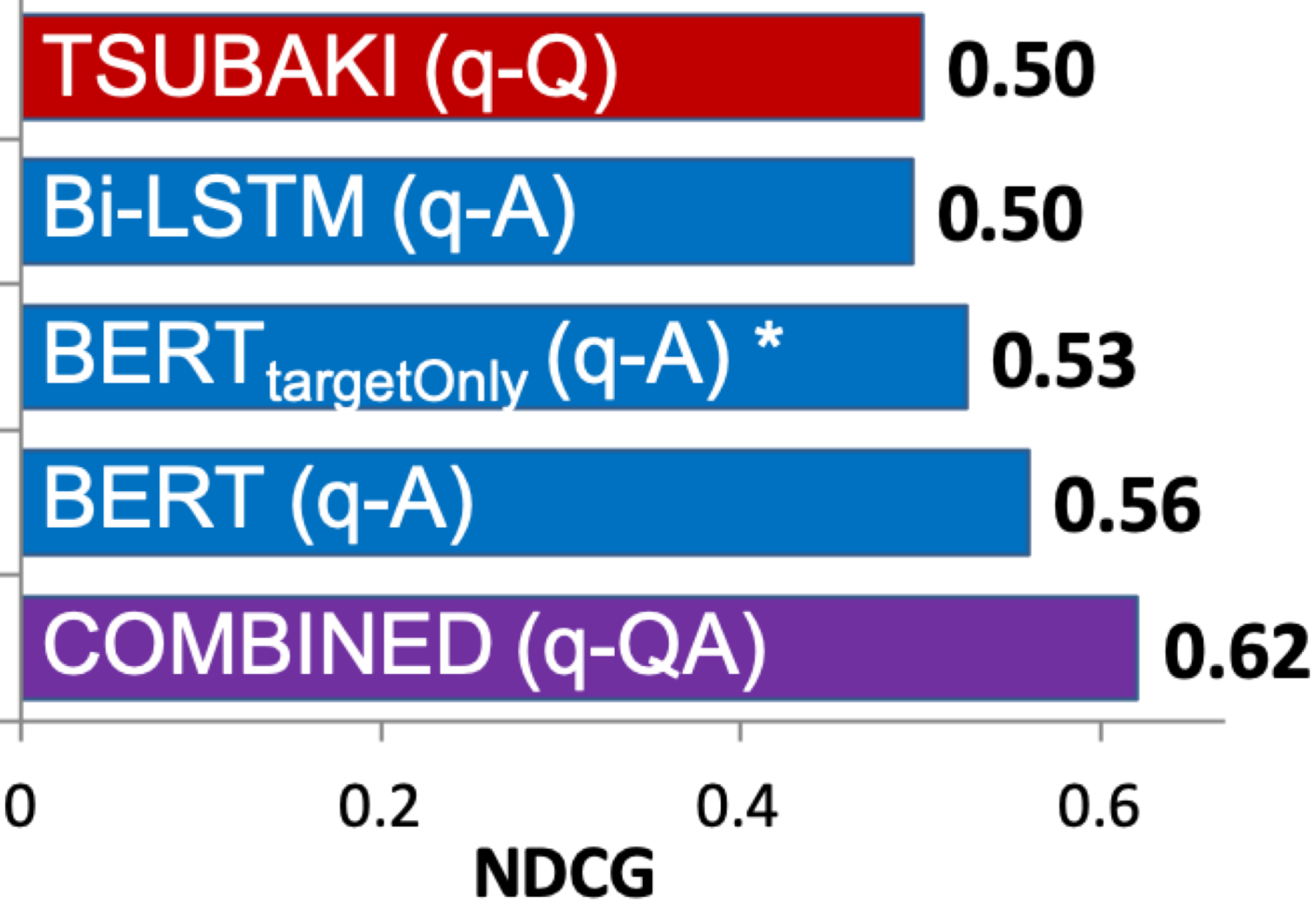
### localgovFAQ dataset

- Collect 784 queries by crowdsourcing and 1786 QAs in FAQ page of Amagasaki City in Japan.
- Each QA pair was manually classified into four categories according to their relevance to query

### Settings

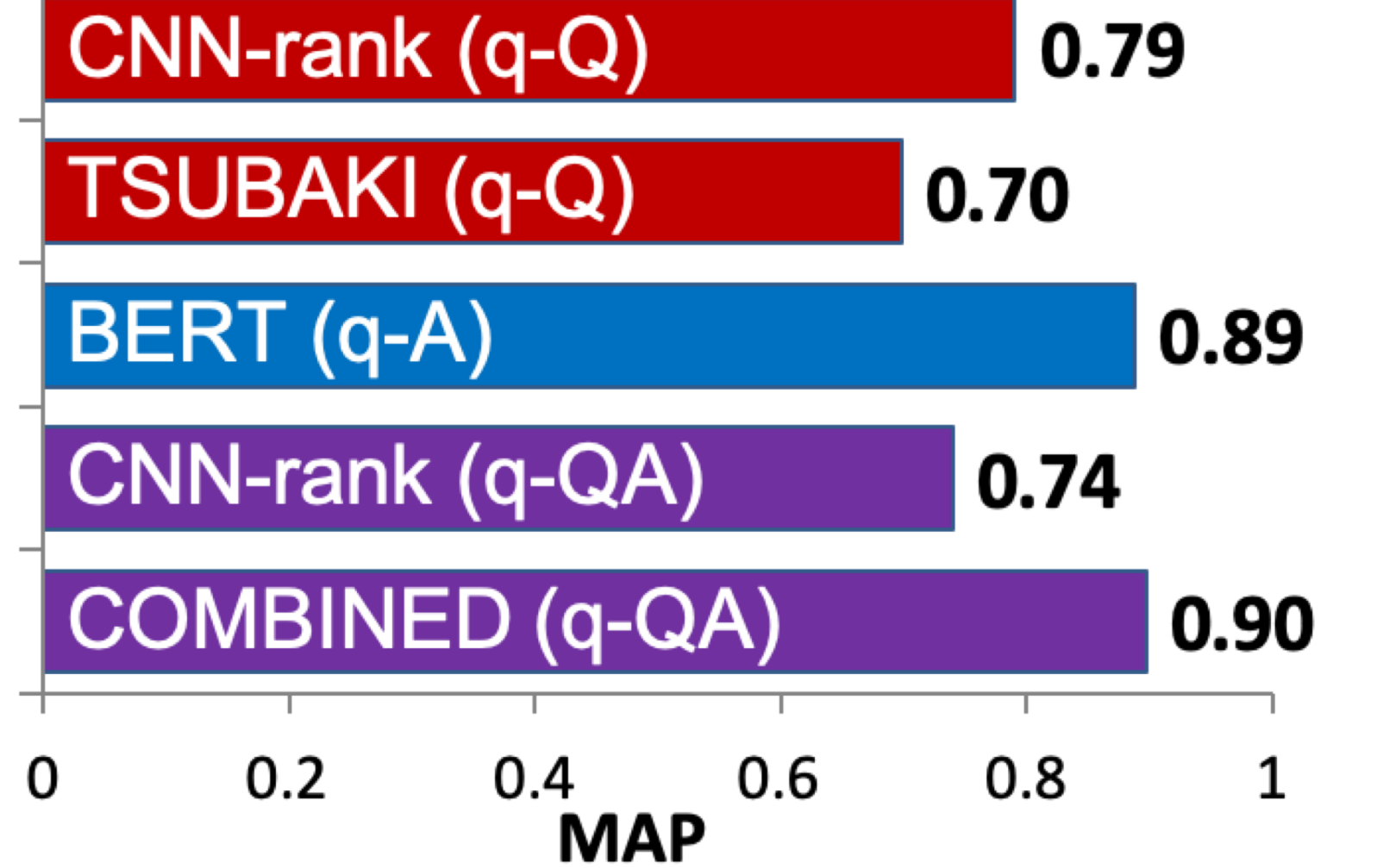
- Use Japanese Wikipedia for pre-training
- Collect 20k QA pairs, using FAQs of 21 Japanese local governments for the fine-tuning.
- Also evaluate in English using Stack Exchange dataset

### localgovFAQ(Japanese)

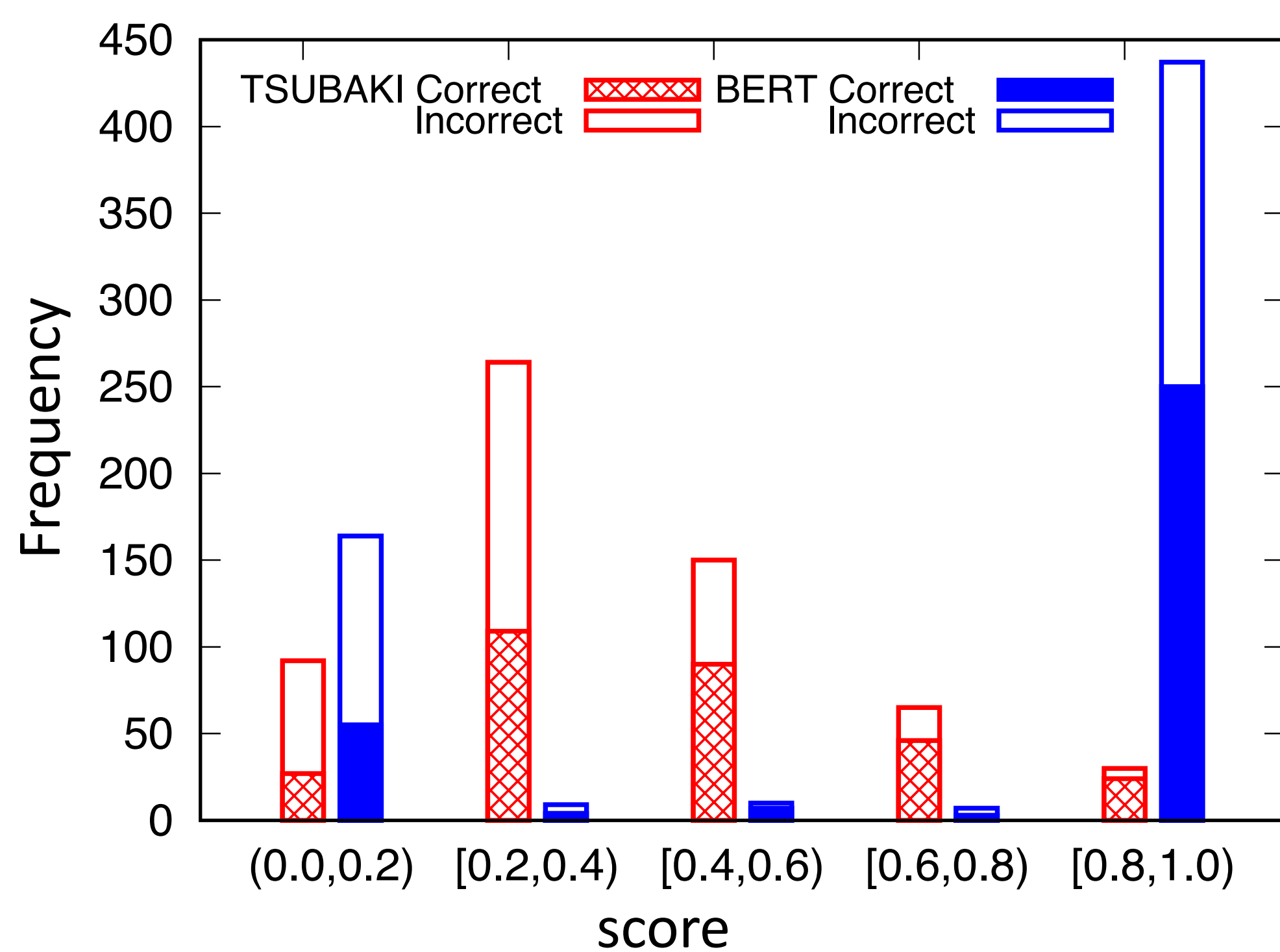


\* BERT<sub>targetOnly</sub>: BERT fine-tuned only with the target FAQ set

### Stack Exchange(English)



## Discussions



- The accuracy of retrieved QA pair with high TSUBAKI score is very high.
- Loose correlation between the accuracy and BERT score.

Query	TSUBAKI	BERT	COMBINED
Is there a consultation desk for workplace harassment?	Q: I'd like to have a career counseling A: Consultation place: Amagasaki City, ...	Q: Can we get labor counseling by a lawyer? A: On specialized and sophisticated labor issues such as wages...	Q: Can we get labor counseling by a lawyer? A: On specialized and sophisticated labor issues such as wages,...
I need the resident certificate, can I take it off hours or on Saturdays and Sundays?	Q: Where and when can I get the resident certificate? A: ... You can get the resident card on weekdays and Saturdays. ...	Q: Can a representative obtain a resident certificate with a personal number? A: The resident certificate with a personal number is ...	Q: Where and when can I get the resident certificate? A: ... You can get the resident card on weekdays and Saturdays. ...
Is there any place where we can use for practicing instruments?	Q: Where is the location of the polling place... A: There are three polling stations before the date in the city. ...	Q: Please tell me about the City boys music corps. A: "The City Boys Music Club" includes a choir corps, a brass band, ...	Q: Please tell me about the City boys music corps. A: "The City Boys Music Club" includes a choir corps, a brass band, ...

Correct Q, "Information on the facility of the youth center, hours of use, and closed day"